
VOICE ANALYZER/SYNTHESIZER

Publication Number: 01-028700 (JP 1028700 A) , January 31, 1989

Inventors:

- YATO TAKASHI

Applicants

- OKI ELECTRIC IND CO LTD (A Japanese Company or Corporation), JP (Japan)

Application Number: 62-184421 (JP 87184421) , July 23, 1987

International Class (IPC Edition 4):

- G10L-007/04

JAPIO Class:

- 42.5 (ELECTRONICS--- Equipment)

JAPIO Keywords:

- R108 (INFORMATION PROCESSING--- Speech Recognition & Synthesis)

JAPIO

© 2005 Japan Patent Information Organization. All rights reserved.

Dialog® File Number 347 Accession Number 2731100

This Page Blank (uspto)

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A)

昭64-28700

⑬ Int. Cl.

G 10 L 7/04

識別記号

庁内整理番号

G-8622-5D

⑭ 公開 昭和64年(1989)1月31日

審査請求 未請求 発明の数 1 (全10頁)

⑮ 発明の名称 音声分析合成器

⑯ 特 願 昭62-184421

⑰ 出 願 昭62(1987)7月23日

⑱ 発 明 者 矢 頭 隆 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内
⑲ 出 願 人 沖電気工業株式会社 東京都港区虎ノ門1丁目7番12号
⑳ 代 理 人 弁理士 大 垣 孝

明 細 書

1. 発明の名称 音声分析合成器

2. 特許請求の範囲

(1) 音声信号の周波数帯域を複数の帯域に分割して得られた各分割チャネル信号を個別に符号化して出力する符号化器と、符号化された分割チャネル信号を受信して合成する復号化器とを含む帯域分割型音声分析合成器において、

前記分割チャネルのうち低周波数域のチャネルに対してはチャネル信号を1サンプル毎に符号化、復号化する符号化器および復号化器を備え、

高周波数域のチャネルに対してはチャネル信号の一定時間区間(フレーム)毎の平均振幅を算出する平均振幅算出回路と、

平均振幅を符号化する符号化器と、

符号化された平均振幅を復号する復号化器と、

復号化されたフレーム平均振幅を1サンプル毎に補間する振幅補間回路と、

符号化されないチャネル信号を代用する雑音発生回路と、

さらに前記雑音発生回路から発生した雑音信号を前記振幅補間回路で生成した1サンプル毎の振幅情報に乗じる乗算器とを備えることを特徴とする音声分析合成器。

3. 発明の詳細な説明

(産業上の利用分野)

この発明は音声分析合成方法及びその装置、特に音声の符号化に関するものである。

(従来の技術)

従来、この種の技術としてザ・ベル・システム・テクニカル・ジャーナル(The Bell System Technical Journal)、55[8](1976-10)(米) P.1069-1085 に記載される帯域分割型音声分析合成方式(Sub-Band Coding 方式とも呼ばれ、以降SBC方式と略す)が知られている。このSBC方式は第3図に示されるように音声信号の周波数帯域を複数(通常4~8)の帯域(図中①、②、③及び④で示す。)に分割し、各分割チャネルの

出力を別々に符号化、復号化する方式である。

第4図にこのSBC方式の基本的な回路構成を示す。また、第5図(A)～(E)は第4図の回路の動作を説明するための図である。以下、第4図、第5図(A)～(E)を用いてSBC方式の動作を説明する。

まず、分析器の動作は次の通りである。マイク(図示せず)等から入力されたアナログ音声信号は、ローパスフィルタ(図示せず)に入力されて所定のサンプリング周波数の $1/2$ 以上の周波数成分を除去された後、A/D変換器(図示せず)で所定のサンプリング周波数においてアナログ信号からデジタル信号 $S_{(n)}$ に変換される。ここで n はサンプル番号である。このデジタル化された入力信号 $S_{(n)}$ はバンドパスフィルタ50に入力され、第5図(A)に示す如く特定の帯域成分(ここでは $W_{1s}-W_{2s}$)が抽出される。次にこのバンドパスフィルタ50の出力信号は乗算器51において第5図(B)に示した W_{1s} なる周波数をもったコサイン波(c o s波)と乗算されることによ

う、この補間部56からの出力信号は、乗算器57において第5図(D)に示した W_{1s} なる周波数をもったc o s波と乗算されることにより復調され、第5図(E)に示した如く基底帯域($0-W_s$)から再びもとの周波数帯域($W_{1s}-W_{2s}$)に戻された後、バンドパスフィルタ58によって信号中の($W_{1s}-W_{2s}$)以外の帯域の成分を除去する。

このようにして、合成器から信号 $S_{s(n)}$ が出力される。

上記一連の処理を各分割帯域(チャンネル)毎にそれぞれ行い、最後に全チャンネルの出力を加算して出力音声信号を得る。

以上がSBC方式の基本的な動作内容であるが、第4図の回路構成を直接装置化することはあまりなく、回路量を削減するためにバンドパスフィルタ50、58を用いない第6図のような構成のSBC方式も提案されている。

次に、この第6図の回路の動作を説明する。

まず、分析器において、デジタル化された入

りc o s変調が施され、第5図(C)の如く($0-W_s$)の基底帯域にシフトされる。このとき生じる $2W_{1s}$ 以上の不要な周波数成分 $R_s(\omega)$ (例えば、第5図(C)で点線で示した成分)をローパスフィルタ52によって除去する。このようにして得られる信号 $r_{s(n)}$ は W_s 以下の周波数成分しか必要としないものであるから、 $2W_s$ のサンプリング周波数でサンプリングすれば必要かつ十分な情報が保たれる。このためにダウンサンプリング部53によって必要以上に高いサンプリング周波数を $2W_s$ に落としてダウンサンプリングを行い、このダウンサンプリングした信号を符号器54で符号化し、符号化された信号を合成器へ伝送する。

次に、合成器において分析器と全く逆の処理を行うことにより、分析器から送られてきた信号を復号する。すなわち、符号化された信号を復号器55によって復号した後、補間部56によって分析器でダウンサンプリングされた信号を元のサンプリング周波数に戻すためにアップサンプリングを行

力信号 $S_{s(n)}$ は複素信号 $e^{j\omega_s n}$ [ここで $\omega_s = (W_{1s} + W_{2s})/2$]にて複素変調される。この複素変調は、乗算器61aによるc o s変調(変調波は $\cos \omega_s n$)、乗算器61bによるサイン(s i n)変調(変調波は $\sin \omega_s n$)により行われる。乗算器61a、61bの出力は帯域幅($0-\omega_s/2$)のローパスフィルタ62a、62bにそれぞれ入力されフィルタリングされる。このようにして、ローパスフィルタ62aからは複素信号 $a_{s(n)} + j b_{s(n)}$ の実部 $a_{s(n)}$ が、ローパスフィルタ62bからは複素信号 $a_{s(n)} + j b_{s(n)}$ の虚部 $b_{s(n)}$ がそれぞれ出力される。各信号 $a_{s(n)}$ 、 $b_{s(n)}$ はそれぞれダウンサンプリング部63a、63bによって周波数 W_s にダウンサンプリングされた後、符号器64によって符号化され、合成器側へ伝送される。合成器においては符号化された信号は復号器65によって復号された後、補間部66a、66bによって元のサンプリング周波数に戻され、次に帯域幅($0-\omega_s/2$)のローパスフィルタ67a、67bを通してフィルタリングさ

れた後、乗算器88aによる \cos 波との乗算、乗算器88bによる \sin 波との乗算によって復調され、さらに加算器88で信号の \cos 成分と \sin 成分とが加算され、当該分割帯域の信号が合成される。

上記一連の処理を各分割帯域（チャンネル）毎にそれぞれ行い、最後に全チャンネルの出力を加算して出力音声信号を得る。

以上がSBC方式の動作原理であるが、この方式は音声信号そのものを符号化する方式に比べ以下のような特長がある。

各チャンネルの量子化誤差は白色雑音に近く、周波数スペクトル上の全域に広がるが、そのうち各チャンネルの帯域内の雑音だけしか各チャンネルには落ちてこないため、量子化雑音を軽減出来る。また、各チャンネルの量子化誤差はその周波数帯域内の信号のみに関係し、音声のように低周波成分が大きく、高周波成分が小さい信号においては周波数の高い帯域のチャンネルでの誤差は信号全体から見れば僅かな誤差にしかない。さらに、音声

H \times z)では3ビット、チャンネル番号3（帯域：1800～2800H \times z)では2ビット及びチャンネル番号4（帯域：2800～3400H \times z)では2ビットである。先に述べたように聴覚的重要度の高い低周波領域に多く、また音質性、量子化誤差に及ぼす影響の少ない高い周波数領域になるほど少ない情報量の割当てとなっている。

ところで、当然のなりゆきとして高品質の合成音をさらに少ない情報量で実現したいという要求がある。SBC方式では帯域分割された各チャンネルの出力信号をサンプル毎に第7図(A)に示されるごとくの量子化ビット数で符号化するが、この出願の発明者等の実験によればこのビット配分は第7図(B)に示すようなビット配分で十分な品質が得られることが分っている。この場合、第7図(A)の場合と異なるのはチャンネル番号2での割り当てビット数が2であることであり、従って、チャンネル番号2の周波数帯域で2ビットの量子化ビット数で十分な品質であればさらに聴覚的重要度の低い2kH \times z付近及びそれ以上の高い周

信号のうち高い周波数の成分は雑音成分が主であり、この帯域での誤差は聴覚上あまり影響しない。

従って、このような性質を考慮して帯域の分割方法や各チャンネルの信号に与える量子化ビット数を設定することにより、音声信号を直接符号化する方式に比べ、約1/2程度の情報量で実現出来る。すなわち、8kH \times zでサンプリングされたPCM音声に対し、これを直接、例えばADPCM符号化した場合、約30Kビット/秒程度の情報量が必要であるが、SBCでは聴覚上ほぼ同品質の合成音が16Kビット/秒前後の情報量で得ることが出来る。

(発明が解決しようとする問題点)

第7図(A)はチャンネルの出力信号を符号化する際のチャンネル毎の量子化ビット数の一般的な割当て例を示す。この図から理解出来るように、チャンネル番号1（帯域：200～1000H \times z)では4ビット、チャンネル番号2（帯域：1000～1800

波数ではもっと粗い量子化を行っても全体の品質に及ぼす影響は少ないと考えられる。事実スペクトル構造上でも2kH \times z以上の周波数域では、それ以下の周波数域に比べ成分も少なく、また音質上の特性を表わすスペクトルのピーク（ホルマント）も低域に比べ明確なものが少ない。上述した点を考慮すればこの領域では思いきった情報量削減が可能と考えられるが、通常のサンプルバイサンプルの符号化方法では情報量の削減に限界があり、高圧縮化の実現が困難であった。

この発明の目的は、上述した従来の問題点を除去し、情報圧縮を十分に行って少ない情報量で良質な合成音が得られるように構成した音声分析合成器を提供することにある。

(問題点を解決するための手段)

この目的の達成を図るため、この発明のSBC方式音声分析合成器においては、帯域分割された各チャンネルの出力信号に対する符号化器として聴覚的に重要な低周波数帯域のチャンネルに対しては

A D P C M、A P C Mなどの1サンプル毎に定まった量子化ビット数にて符号化する符号化器及び復号化器を具えている。

さらに合成音の品質に及ぼす影響の少ない高周波数帯域のチャンネルに対しては一定フレーム毎にチャンネル信号の平均振幅を算出する平均振幅算出器と、算出された平均振幅を符号化する対数圧伸符号化器と、符号化された平均振幅を復号化する対数伸長器と、復号化された平均振幅をサンプル毎に補間する振幅補間器と、平坦なスペクトルを持った信号を発生するノイズ発生器と、発生したノイズと振幅補間された振幅情報とを乗じてチャンネル信号とする乗算器とを設けたものである。

(作用)

上述したこの発明によれば、2 k H z 付近あるいは2 k H z 以上の周波数のチャンネル成分に対してサンプルバイサンプルの符号化を行わず、一定時間区間のフレーム(10~20 m s e c)毎にチャンネル信号の平均振幅を算出し、算出された平均振

幅の構成を示すブロック図であり、10は入力端子、11a₁~11a₄及び11b₁~11b₄は乗算器、12a₁~12a₄及び12b₁~12b₄はローパスフィルタ、13a₁~13a₄及び13b₁~13b₄はR:1のダウンサンプリング部、14a₁~14a₄及び14b₁~14b₄は符号化部で、これらは分析側の構成部分であって全チャンネルとも第6図に示した分析器の構成に対応する。さらに、合成側の装置部分も第6図の合成器の構成と対応して構成しており、第1図において15a₁~15a₄及び15b₁~15b₄は復号化部、16a₁~16a₄及び16b₁~16b₄は1:Rの補間器、17a₁~17a₄及び17b₁~17b₄はローパスフィルタ、18a₁~18a₄及び18b₁~18b₄は乗算器、19は加算器、20は出力端子である。さらに110及び111はそれぞれマルチプレクサ及びデマルチプレクサである。そして全体としての構成は、従来と同様に、音声信号の周波数帯域を複数の帯域に分割し、各分割チャンネル信号を別個に符号化し合成するようになしてある。

幅情報のみを符号化して合成部へ送出し、合成部では白色雑音をサンプル信号に代用して、送られてきたフレーム振幅情報を白色雑音に乗じてチャンネルの平均振幅のみを保存したチャンネル信号を再生する構成となっているので、大幅な情報量削減を図ることが出来る。

また、この発明によれば、フレーム毎に送られてくる平均振幅情報をサンプル毎に補間するので、フレーム間での電力の不連続をなくし、自然音に近い合成音を得る。

(実施例)

以下、図面を参照してこの発明の実施例につき説明する。

全体構成

初めに全体構成について述べる。第1図は第6図に示した従来のS B C方式音声分析合成器の構成に基づいて構成された、この発明の実施例である4チャンネルS B C方式音声分析合成器の全チャ

各分割チャンネル内での動作

第1図に示す合成分析器において、チャンネル1~チャンネル4の各分割チャンネル内での動作は、符号化器の内部処理を除いて、全く同じである。また同一チャンネル内でも構成部分11a₁~11a₄(但し、kはk番目のチャンネルを表わしている。この実施例ではk=1~4である。)のcos成分に対するブロックと11b₁~11b₄までのsin成分に対するブロックとでは変調波がcosとsinで異なるだけで動作は全く同じであるため、ここでは1つのチャンネル例えば代表してk=1のチャンネルのcos成分に対する動作について以下に説明する。

まず、入力端子10よりディジタル化された音声信号が入力されると、その信号に対しチャンネルの中心周波数と同じ周波数を持ったcos波形(cos ω₁ t この実施例ではcos ω₁ t)を乗じ変調を行う。cos変調された音声信号はω₁すなわちω₁の1/2の帯域をもったローパスフィルタ12a₁に通され、このチャンネルのcos

成分の出力 $a_i(n)$ が抽出される。次にローパスフィルタ12a_iの出力 $a_i(n)$ はダウンサンプリング部13a_iにおいて(チャンネルの帯域幅)/(元の信号のサンプリング周波数)のサンプルにダウンサンプリングされ、その結果 $a_i(SR)$ を符号化器14a_iによって符号化し、符号化結果 $A_i(SR)$ をマルチプレクサ110を介して合成側へ送出する。

分析側ではデマルチプレクサ111を介して分析側より送られてきた符号化結果 $A_i(SR)$ を復号化器15a_iにより復号し $a'_i(SR)$ を得る。 $a'_i(SR)$ は補間器16a_iによって元のサンプリング周期に戻され、ローパスフィルタ17a_iを通り、さらに乗算器18a_iにおいて $\cos \omega_i n$ を乗せられて再び元の周波数帯域に復元される。

以上の処理を \sin 成分、さらに $k=2, 3$ 及び4等の他チャンネルも同様にして行い、最後に全チャンネルの出力結果を加算器19により加算し、合成結果として出力端子20から出力する。

部のブロック図である。

まず、第2図(A)を参照してAPCM符号化の動作について説明する。第2図(A)において14は符号化部、15は復号化部、21は入力バッファ、22は符号化器である。23~28は符号化器22における量子化の量子化ステップ幅を決定するブロックでこれを量子化ステップ幅決定部29として示す。このブロック29において23は絶対値回路、24は最大値検出回路、25は量子化レベル符号化器、26は対数圧伸テーブル、27は量子化レベル復号化器、28は対数伸長テーブルである。30、31はそれぞれマルチプレクサ、デマルチプレクサで、32は復号化器、33は量子化レベル復号化器、34は対数伸長テーブルである。

以下、動作を述べる。入力信号は符号化器22によってサンプル毎に符号化される訳であるが、この場合量子化ステップ幅はフレーム内の入力信号の最大値が量子化のダイナミックレンジに等しくなるように量子化ステップ幅決定部29で決定される。21は1フレーム間の量子化ステップ幅が算出

符号化器、復号化器部分の説明

次に、この発明の要旨である符号化器、復号化器部分の動作についてさらに詳細に説明する。

この実施例では、一例として、チャンネル1からチャンネル3までの符号化器14a₁~14a₃、14b₁~14b₃、復号化器15a₁~15a₃、15b₁~15b₃にはサンプルバイサンプルの符号化方式APCMを用いる。又、粗い符号化が許容できる高周波数帯域のチャンネル4の符号化器14a₄~14b₄及び復号化器15a₄~15b₄に先に述べたフレーム平均振幅を符号化する方法を用いている。

<APCM符号化・復号化部>

第2図(A)はチャンネル1~3で採用されるAPCM符号化・復号化部すなわち、例えば第1図に200で示す点線枠の部分の構成を具体的に示すAPCM符号化・復号化部のブロック図である。第2図(B)は平均振幅符号化及び再生部すなわち例えば第1図に400で示す点線枠の部分の構成を具体的に示す平均振幅符号化・復号化再生

されるまで1フレーム分の入力信号を蓄えるバッファであり、入力信号はまず、その振幅絶対値を絶対値回路23で算出し、さらにフレーム内でのその最大値を最大値検出回路24で求める。当然のことながら符号化で用いた量子化ステップ幅は復号化器32でも用いるため、ステップ幅を決定するフレーム最大振幅の情報を合成側にも送る必要がある。従って、求まったフレーム最大値幅をここでは量子化レベル符号化器25において対数圧伸してビット数を削減し、合成側へ送出する。対数圧伸は対数圧伸テーブル26を用いて行われる。対数圧伸符号化されたフレーム振幅最大値は対数伸長テーブル28を用いて量子化レベル復号化器27によって復号されるが、この場合、対数伸長テーブル27には、符号化器22の量子化ビット数を P とすると、対数圧伸テーブルの値に対して $1/2^{P-1}$ の値が格納されており、フレーム最大振幅の復号に際して元の値の $1/2^{P-1}$ 倍が出力されて直接符号化器22における量子化ステップ幅となる。この量子化ステップ幅を用いて符号化器22では入力

信号バッファ21に蓄えられている1フレーム分の入力信号を符号化し、復号化部15へ送出する。復号化部15では送られてきた量子化ステップ幅符号化情報を対数伸長テーブル34を用いて量子化レベル復号化部33によって復号し復号化部32に与える。この量子化レベル復号化部33と対数伸長テーブル34は符号化部14の量子化レベル復号化部27及び対数伸長テーブル28と全く同一のものである。復号化部32ではこうして与えられた量子化ステップ幅とを用いて1フレーム分の符号化データの復号を行う。

<平均振幅符号化・再生部>

次に第2図(B)の平均振幅符号化部・再生部について説明する。第2図(B)において41は絶対値回路、42は平均振幅算出器、43は平均振幅符号化器、44は対数圧伸テーブル、45は平均振幅復号化器、46は対数伸長テーブル、47は振幅補間回路、48は白色雑音発生器、49は乗算器である。

以下、動作について説明する。

の振幅情報とする。これにより品質の良い合成音が得られる。以上のごとく復号されたサンプル毎の振幅情報と雑音発生器48の出力信号を乗算器49にて掛け合わせ当該チャンネルの出力とする。

この発明は上述した実施例にのみ限定されるものではない。上述した実施例においては、第4チャンネルのみに対して上記の処理を施したが、さほど高品質を要求されない用途であれば3チャンネルに対しても同様な処理を行い、さらに情報圧縮を行うことも十分可能である。

(発明の効果)

上述した説明からも明らかなようにこの発明では聴感上あまり重要でない高域のチャンネルに対してはサンプル毎の符号化を行わず、フレーム振幅情報しか送らないため、情報量が大幅に削減出来る。

この発明は少ない情報量で高品質な合成音を得られる。

入力信号は、まず、その振幅絶対値を絶対値回路41で算出する。平均振幅算出器42では算出された振幅絶対値を1フレーム分逐次加算し、平均振幅符号化器43によって対数圧伸符号化する。対数圧伸は対数圧伸テーブル44を用いて行う。こうして得られた平均振幅符号化データは合成側の復号化部15へ送られる。合成側では送られてきた平均振幅符号化データは平均振幅復号化器45において対数伸長テーブル46を用いて対数伸長されるが、符号化側で符号化した値は平均振幅ではなく、フレーム内の振幅絶対値の総和であるため、対数伸長テーブル46には、対数圧伸テーブル44に対して $1/(1\text{フレームのサンプル数})$ した値を格納しておく。ところで、第2図(B)に示す平均振幅符号化・再生部の回路では1フレームに1つの平均振幅情報しか合成側に伝送しないため、これをそのまま1フレーム内の全てのサンプルに対して用いるとフレーム間での電力の不連続が生じてしまう。これがため、ここでは振幅補間回路47によって1サンプル毎に振幅補間を行いサンプル毎

4. 図面の簡単な説明

第1図はこの発明の音声分析合成器の全体構成の実施例を示すブロック図。

第2図(A)はAPCM符号化・復号化部のブロック図。

第2図(B)は平均振幅符号化・再生部のブロック図。

第3図はSBC方式の説明図。

第4図は従来のSBC方式音声分析合成器の構成を示すブロック図。

第5図は第4図の装置の動作を説明するための図。

第6図は他の従来のSBC方式音声分析合成器の構成を示すブロック図。

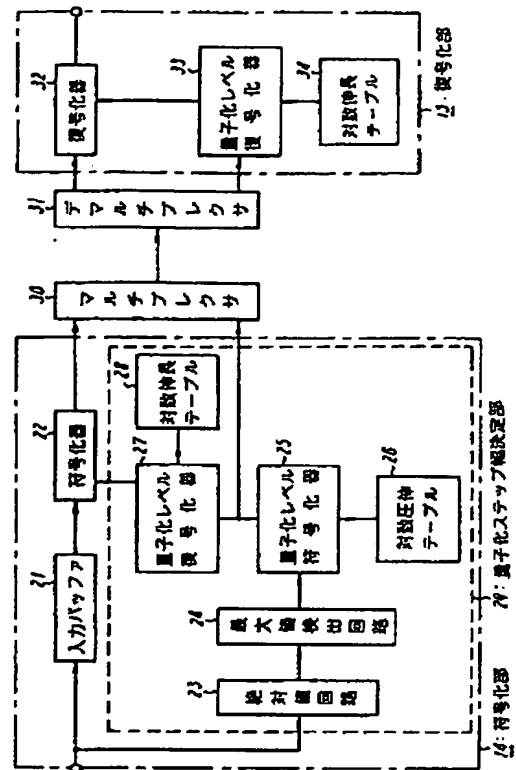
第7図は周波数帯域と割り当てビット数との関係を示す図である。

10—入力端子

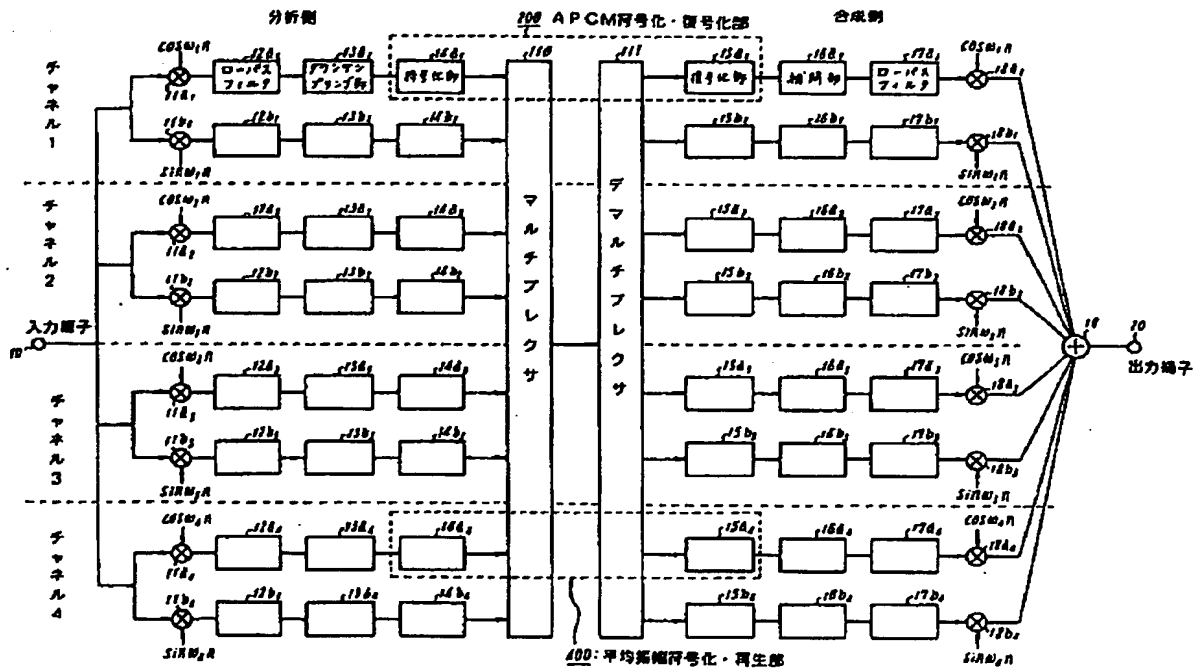
11a₁~11b₂、18a₁~18b₂、49—乗算器

12a₁~12b₂、17a₁~17b₂—ローパスフィルタ

- 13a, ~ 13b, ... ダウンサンプリング部
 14, 14a, ~ 14b, ... 符号化部
 15, 15a, ~ 15b, ... 復号化部
 16a, ~ 16b, ... 補間器, 19 - 加算器
 20 - 出力端子, 30, 110 - マルチプレクサ
 31, 111 - デマルチプレクサ
 200 ... APCM符号化・復号化部
 400 ... 平均振幅符号化・再生部
 21 - 入力バッファ, 22 - 符号化器
 23, 41 - 絶対値回路, 24 - 最大値検出回路
 25 - 量子化レベル符号化器
 27, 33 - 量子化レベル復号化器
 28, 34 - 対数伸長テーブル
 29 - 量子化ステップ幅決定部
 32 - 復号化器, 42 - 平均振幅算出器
 43 - 平均振幅符号化器, 44 - 対数圧伸テーブル
 45 - 平均振幅復号化器, 46 - 対数伸長テーブル
 47 - 振幅補間回路, 48 - 白色雑音発生器。

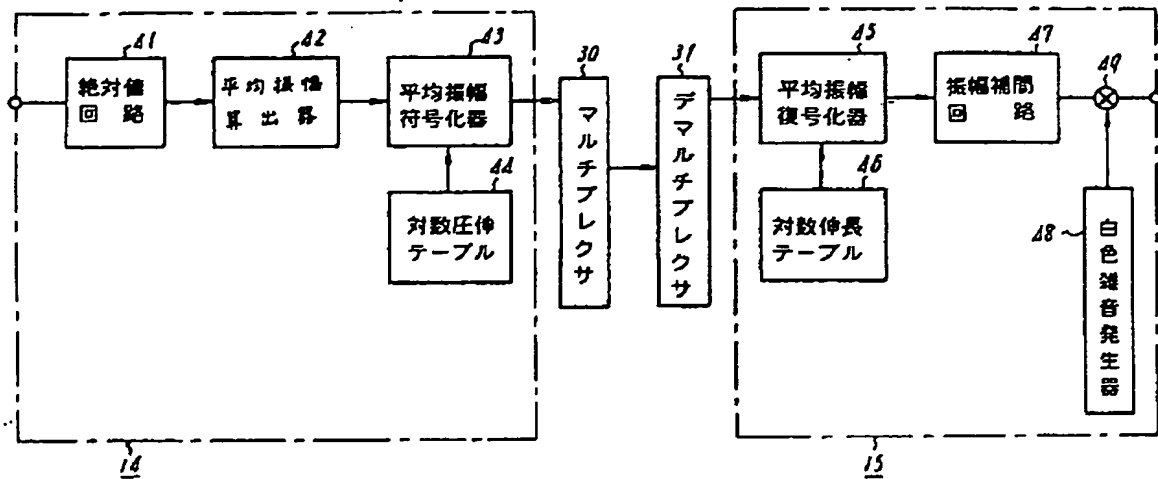


APCM符号化・復号化部のブロック図
 第2図(A)



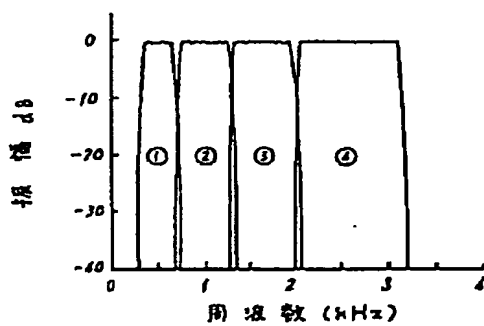
この発明の音声分析合成器の実施例

第1図



平均振幅符号化・再生部のブロック図

第 2 図 (B)



SBC方式の説明図

第 3 図

(A)

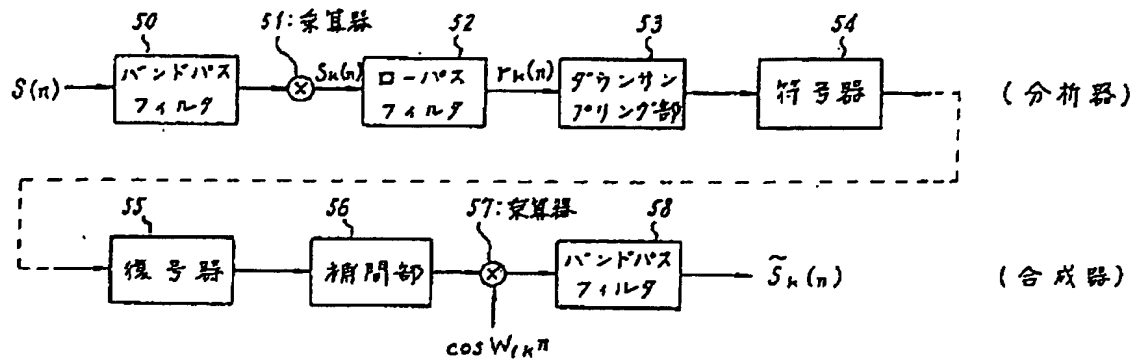
チャンネル番号	帯域 (Hz)	割り当てビット数
1	200 ~ 1000	4
2	1000 ~ 1800	3
3	1800 ~ 2600	2
4	2600 ~ 3400	2

(B)

チャンネル番号	帯域 (Hz)	割り当てビット数
1	200 ~ 1000	4
2	1000 ~ 1800	2
3	1800 ~ 2600	2
4	2600 ~ 3400	2

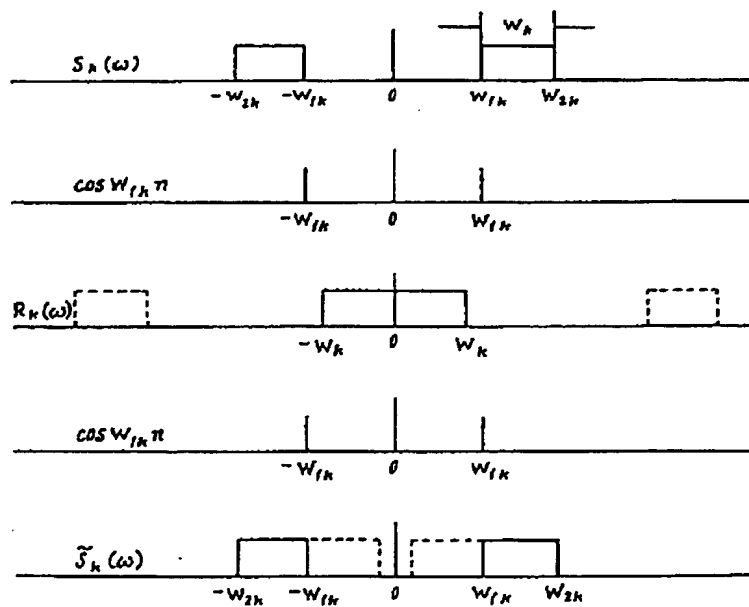
周波数帯域と割り当てビット数との関係

第 7 図



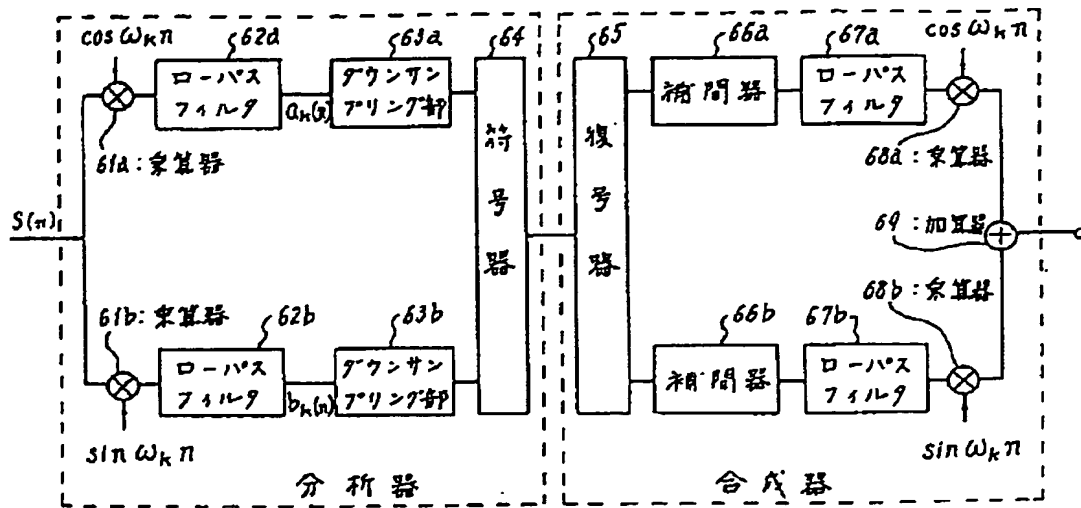
従来の SBC 方式音声分析合成器の構成

第 4 図



第 4 図の装置の動作を説明するための図

第 5 図



他の従来のSBC方式音声分析合成器の構成

第 6 図